Los Alamos
NATIONAL LABORATORY
memorandum

Applied Theoretical & Computational Physics
Transport Methods Group
http://www-xdiv.lanl.gov/XTM/

To/MS: Distribution
From/MS: H. Lichtenstein (XTM)/B226
Phone/FAX: 7–3926/5–5538
Email: hal@lanl.gov
Symbol: XTM:97–122 (U)
Date: May 7, 1997

SUBJECT: Preliminary to X-Division Research Note

"PARAMETRIC INVESTIGATION OF POTENTIAL FALSE LEARNING IN ADAPTIVE MONTE CARLO TRANSPORT"

Abstract

Having generalized a code ("Code"), which Tom Booth (XTM) had used originally to test his algorithms for adaptive Monte Carlo transport, the subject effort utilized the generalized Code to investigate potential false learning (FL) situations. The generalized Code allows specifying multiple spatial domains (*i.e., line-intervals*, spanning *multiple mean-free-paths* [mfp]), for the so-called *tri-directional problem*.

Attention was focused on the efficacy of a strategy to preclude FL, which was devised by Booth. A viable basis for diagnosing the presence of FL may be a comparison between theoretical and computed values of quantities related to local behavior. Such comparisons are based on the recognition that, although the global solution is unknown, local behavior *is* known.

The search to uncover FL was based on varying several parameters of the Code designed to solve the generic tri-directional problem. The principal finding was the *absence* of any evidence for FL.

Preface

This memorandum, having a small distribution, is intended to serve as a draft for a wider-distribution research note. Members of the present distribution are invited to respond with comments and suggestions for modifying this memorandum to produce the subsequent research note.

1. Introduction

This research is a continuing effort, addressing the issue of potential false learning (FL) in adaptive Monte Carlo transport. The efforts that enabled the investigation described herein have been documented previously: Booth (XTM) established the theoretical and algorithmic foundation;[1, 2] and I contributed supporting empirical findings,[3] and a generalization of Booth's test-bed code ("Code").[4, 5]

In the following sections, I begin with a review of Booth's strategy to preclude FL,[2] preserving his enumeration of equations to facilitate cross referencing. Then I describe my recent efforts to address some of the questions that remained at the issuance of Booth's comprehensive research note.[2]

2. Background

Booth's theory[2] incorporates some "insurance against FL," based on the recognition that, although the global solution is unknown, local behavior *is* known. This "insurance policy" is summarized below, omitting the derivations given in [2].

The biased probability of a particle colliding in dy after traveling forward a distance y from x is given by

$$c(x, x+y) = \frac{e^{-\sigma y}}{N(x)}(\sigma N(x+y) - N'(x+y)) \tag{23}$$

where σ is the total cross section and $N(x)$ is the forward importance (expected score for particle moving forward) at x.

The associated weight multiplication for a particle colliding at $x+y$ moving forward is

$$w_c(x, x+y) = \frac{\sigma N(x)}{\sigma N(x+y) - N'(x+y)} \tag{24}$$

Similarly, for a particle traveling backward

$$c(x, x-y) = \frac{e^{-\sigma y}}{L(x)}(\sigma L(x-y) + L'(x-y)) \tag{28}$$

and

$$w_c(x, x-y) = \frac{\sigma L(x)}{\sigma L(x-y) + L'(x-y)} \tag{29}$$

where $L(x)$ is the backward importance at x.

Note that the collision density $c(x, x+y)$ (Eq. 23) goes to zero when the expression

$$\sigma N(x+y) - N'(x+y)$$

vanishes, with the corresponding weight multiplication going to infinity (Eq. 24).

Similarly, $c(x, x-y)$ (Eq. 28) goes to zero when the expression

$$\sigma L(x-y) + L'(x-y)$$

vanishes, with the corresponding weight multiplication going to infinity (Eq. 29). In order to have positive bounded weights it is necessary to bound $N'(x)$ and $L'(x)$.

Before addressing the need for bounded derivatives of the importance, it is also pertinent to observe that escapes (*i.e.*, reflections) at the source boundary have unbounded weight (as originally posed), *viz.*

$$w_u(x, x-y) = \frac{L(x)}{L(x-y)} \tag{17}$$

where $L(x-y)$ vanishes for $x-y=0$. This source of trouble can be avoided if the score function is modified.

The effects of adding δ to any history termination are: (a) score δ if the particle is absorbed, or if it escapes at x=0 (reflection); and, (b) score 1+δ if the particle escapes at x=thickness (penetration). Hence, this new score function adds exactly δ to the score of each particle, and the mean penetration is just the new mean score minus δ.

Given such a score function, Booth derives the differential form of the integral importance equations (which latter he derived in reference [6]), *viz.*

$$N'(x) - (\sigma - \sigma_s f)N(x) = -\sigma_s r L(x) - [\sigma - \sigma_s]\delta \tag{57}$$

and

$$-L'(x) - (\sigma - \sigma_s f)L(x) = -\sigma_s r N(x) - [\sigma - \sigma_s]\delta \tag{60}$$

In Eq. 57, $L(x) \geq 0$, and in Eq. 60, $N(x) \geq 0$. This allows the derivatives to be bounded, *viz.*

$$N'(x) \leq N(x)(\sigma - \sigma_s f) - [\sigma - \sigma_s]\delta \tag{61}$$

and

$$-L'(x) \leq L(x)(\sigma - \sigma_s f) - [\sigma - \sigma_s]\delta \tag{62}$$

Requiring the conditions of Eqs. 61 and 62 everywhere ensures that the collision densities in Eqs. 23 and 28 are ≥ 0, and that the denominators in Eqs. 24 and 29 are bounded away from zero. Thus, unbounded weight multiplications are precluded. The δ-modified score function and the bounded derivatives of the importance comprise the basis of Booth's insurance against FL.

Testing the efficacy of these insurance measures required the ability for the transport to cover more than a mean-free-path [mfp] on the line, which was the approximate limitation of Booth's original Code. My generalization of the Code [4, 5] enabled such testing, the results of which are presented in the sections that follow.

As a final background note, Booth alluded to the possible existence of "a bug in the coding of the false convergence protection" ([2], p.24). In the course of my testing, I did, in fact, encounter and correct a bug in the routines that incorporate the conditions of Eqs. 61 and 62.

3. Parametric Studies

I proceeded to drive the generalized Code [5] in an attempt to demonstrate FL. Two of the obvious parameters that could be examined are "thickness" (*i.e.*, total line-length [mfp]) and δ. Other parameters that characterize this generic problem are: "kmax" (*i.e.*, number of segments for importance function bookkeeping); number of histories per iteration (*i.e.*, batch size), where total number of iterations is fixed at 10; and partition of probability for forward/backward scattering.

Tables I–V below summarize the computed results. The **theoretical** values were computed by a code based on the analytic solution.[7] Unless otherwise stated, the partition between forward and backward scattering probability was equal (transverse scattering is ignored, since it is essentially a *no-op*); batch size was 5000 histories per iteration (with a fixed total of 10 iterations); and 10 line segments were used for computing the importance (forward and backward) at the segment nodes.

4. Results and Discussion

TABLE I
Results for Thickness=1 mfp Compared with Theory

δ	mean	rel. error	rel. $\lvert mean - theory \rvert$	δ / mean
1.E-00	.48200365430	$\sim$ 0.0	$\sim$ 0.0	2.E-00
1.E-01	.48200365430	4.E-09	$\sim$ 0.0	2.E-01
1.E-02	.48200365430	6.E-09	$\sim$ 0.0	2.E-02
1.E-03	.48200365430	7.E-09	$\sim$ 0.0	2.E-03
1.E-04	.48200365430	$\sim$ 0.0	$\sim$ 0.0	2.E-04
1.E-05	.48200365430	9.E-10	$\sim$ 0.0	2.E-05
1.E-06	.48200365430	3.E-09	$\sim$ 0.0	2.E-06
1.E-07	.48200365430	$\sim$ 0.0	$\sim$ 0.0	2.E-07
1.E-08	.48200365430	$\sim$ 0.0	$\sim$ 0.0	2.E-08
1.E-09	.48200365430	$\sim$ 0.0	$\sim$ 0.0	2.E-09
1.E-10	.48200365430	$\sim$ 0.0	$\sim$ 0.0	2.E-10
1.E-11	.48200365430	4.E-09	$\sim$ 0.0	2.E-11
	.48200365430 (theory)			

Booth had speculated that a useful value of δ may be on the order of the mean.[7] But the results in Table I fail to show any sensitivity to δ, over a huge range of δ values. At least for this easily obtained mean value, there is no indication of a preferential link between mean value and value of δ. This observation may be useful to a strategy for generalizing the selection of δ in situations where the mean to be computed is not easily estimated.

TABLE II
Results for Thickness=4 mfp Compared with Theory

δ	mean	rel. error	rel. $\lvert mean - theory \rvert$	δ / mean
1.E-00	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E+01
1.E-01	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E-00
1.E-02	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E-01
1.E-03	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E-02
1.E-04	5.7371735870D-02	5.E-09	$\sim$ 0.0	2.E-03
1.E-05	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E-04
1.E-06	5.7371735870D-02	6.E-09	$\sim$ 0.0	2.E-05
1.E-07	5.7371735870D-02	4.E-09	$\sim$ 0.0	2.E-06
1.E-08	5.7371735870D-02	2.E-09	$\sim$ 0.0	2.E-07
1.E-09	5.7371735870D-02	4.E-09	$\sim$ 0.0	2.E-08
1.E-10	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E-09
1.E-11	5.7371735870D-02	$\sim$ 0.0	$\sim$ 0.0	2.E-10
1.E-12	5.7371735870D-02	4.E-09	$\sim$ 0.0	2.E-11
	5.7371735870D-02 (theory)			

Again, as for the thickness of 1 mfp (Table I), the 4-mfp results in Table II show no sensitivity to the value of δ used.

TABLE III
Results for Thickness=7 mfp Compared with Theory

δ	mean	rel. error	rel. $\lvert mean - theory \rvert$	δ / mean
1.E-00	6.8768457934D-03	4.E-05	3.E-05	1.E+02
1.E-01	6.8767713085D-03	2.E-06	2.E-05	1.E+01
1.E-02	6.8767991806D-03	2.E-04	2.E-05	1.E-00
1.E-03	6.8766511993D-03	6.E-06	2.E-06	1.E-01
1.E-04	6.8767312250D-03	2.E-05	1.E-05	1.E-02
1.E-05	6.8765474893D-03	3.E-05	1.E-05	1.E-03
1.E-06	6.8766639654D-03	1.E-05	4.E-06	1.E-04
1.E-07	6.8768783692D-03	5.E-05	3.E-05	1.E-05
1.E-08	6.8769743653D-03	4.E-05	5.E-05	1.E-06
1.E-09	6.8769743649D-03	4.E-05	5.E-05	1.E-07
1.E-10	6.8769743648D-03	4.E-05	5.E-05	1.E-08
1.E-11	6.8769743648D-03	4.E-05	5.E-05	1.E-09
	6.8766393483E-03 (theory)			

As for the foregoing cases, the Table III results show little sensitivity to δ. But the batch size of 5000 histories per iteration appears inadequate for 7 mfp. The best agreement with theory appears to be for a δ of 0.001. Hence, using this value of δ, I proceeded to examine how convergence for 7 mfp responds to batch size.

TABLE IV
Results for Thickness=7 mfp *vs.* Batch Size
δ=0.001; δ/mean = 0.1

batch	mean	rel. error	rel. $\|mean - theory\|$
5K	6.87665120D-03	6.E-06	2.E-06
10K	6.87664560D-03	1.E-06	9.E-07
1000K	6.87663935D-03	3.E-09	$\sim$ 0.0
100K	6.87663935D-03	$\sim$ 0.0	$\sim$ 0.0
50K	6.87663935D-03	$\sim$ 0.0	$\sim$ 0.0
25K	6.87663935D-03	$\sim$ 0.0	$\sim$ 0.0
	6.87663935E-03 (theory)		

The variation of batch size in the Table IV results is presented in *chronological* order (top-to-bottom) to describe my thought process. Having worker machine "wings" at Booth's and my exclusive disposal, at least for the time being, allowed me to submit a 10 Mega-history job with impunity (so as to drive down the relative error to "convergence" level, albeit by overkill). I did not record how long it took; I submitted the job one Friday evening and it was done the following Monday morning. Subsequently, I backed off to 100K, 50K, and 25K batch size, to find the the approximate minimal number of histories per iteration for convergence.

TABLE V
Results for Thickness=8–10 mfp *vs.* Batch Size
δ =0.001

thickness	batch	mean	rel. error	rel. $\|mean - theory\|$
8 mfp	25K	3.3906517711D-03	$\sim$ 0.0	4.E-10
		3.3906517723E-03 (theory)		
9 mfp	25K	1.6711331777D-03	4.E-4	4.E-04
	50K	1.6715238785D-03	8.E-5	2.E-04
	100K	1.6718237777D-03	$\sim$ 0.0	$\sim$ 0.0
		1.6718237777E-03 (theory)		
10 mfp	100K	8.2432332089D-04	8.E-6	7.E-07
	200K	8.2432385602D-04	1.E-7	6.E-08
	300K	8.2432391920D-04	9.E-9	1.E-08
	400K	8.2432390559D-04	1.E-8	3.E-09
	600K	8.2432384570D-04	3.E-7	8.E-08
	1000K	8.2431850941D-04	8.E-6	7.E-06
	2000K	8.2432413940D-04	2.E-7	3.E-07
		8.2432390783E-04 (theory)		

The results in Table V, for 8–9 mfp, are un-remarkable. But for 10 mfp, it is curious that convergence appears to peak for a batch size of $\sim$ 400K. I am not sure if this is significant or just an anomaly of no consequence.

5. Summary

Adaptive Monte Carlo transport relies on learned information to accelerate convergence to a zero-variance biasing solution. Such an iterative procedure may be vulnerable to false learning (FL). As reported by Booth,[2] he has incorporated in the Code a procedure that may preclude FL. I have also reported[3] some evidence that the *tendency* to inspect all states has the *tendency* to avoid FL. And we have both identified a potential basis for diagnosing the presence of FL, namely, a comparison between theoretical and computed values of quantities related to local behavior.[2, 3] Such comparisons are based on the recognition that, although the global solution is unknown, local behavior *is* known.

The efficacy of Booth's "insurance against FL" has been investigated in this current effort, using the Code that is documented on the WWWeb.[5] With many parameters available for investigation, this study can not claim to be definitive, or even conclusive. **But no evidence of FL has been detected in this effort.**

Of secondary importance is the finding that computation of the problem mean is insensitive to the value of δ used. This finding may help define an appropriate general-purpose algorithm for an adaptive Monte Carlo scoring function.

Finally, the curious behavior of the convergence observed for the 10-mfp case may bear further study.

REFERENCES

[1] T. E. Booth, "Tridirectional Problem Discretization for LDRD," *WWW URL http://www-xdiv.lanl.gov/XTM/projects/mc21/wip/tran/memos/*, Document Booth#1.

[2] T. E. Booth, "Adaptive Monte Carlo Attempts on a Continuous Transport Problem," X-Division Research Note XTM–RN (U) 96–013, to Distribution, *WWW URL http://www-xdiv.lanl.gov/XTM/rn/rnotes.html* (December 13, 1996).

[3] H. Lichtenstein, "Investigation of False Learning in Adaptive Monte Carlo Transport," X-Division Research Note XTM–RN (U) 96–009, to Distribution, *WWW URL http://www-xdiv.lanl.gov/XTM/rn/rnotes.html* (August 15, 1996).

[4] H. Lichtenstein, "Generalization of Booth's Code for Adaptive Monte Carlo Transport," X-Division Research Note XTM–RN (U) 97–009, to Distribution, *WWW URL http://www-xdiv.lanl.gov/XTM/rn/rnotes.html* (March 28, 1997).

[5] H. Lichtenstein, "Generalized Code: Source Listing," *WWW URL http://www-xdiv.lanl.gov/XTM/projects/mc21/wip/tran/memos/*, Document Lichtenstein#7a.

[6] T. E. Booth, *Nucl. Sci. Eng.*, **122**, 79 (1996).

[7] Tom Booth, private communication.

HL:hal

Distribution:
P. D. Soran, XTM, MS B226
D. G. Shirk, XTM, MS B226
W. T. Urban, XTM, MS B226
T. E. Booth, XTM, MS B226
R. A. Forster, XTM, MS B226
H. G. Hughes, XTM, MS B226
R. C. Little, XTM, MS B226
R. E. Prael, XTM, MS B226
H. Lichtenstein, XTM, MS B226
XTM file